Breaking the Cancer Genome Code for Patient Care

Sie, D.L.S.

2017

document version
Publisher's PDF, also known as Version of record

Chapter 1

Molecular cytogenomic analysis of solid tumors by next generation sequencing

Daoud Sie
Daniëlle Heideman
Erik A Sistermans
Donna G Albertson
Bauke Ylstra

Adapted from
Sie, D., Sistermans, E., Heideman, D., & Ylstra, B. (2017). Cytogenomics of Solid Tumors by Next-Generation Sequencing: A Clinical Perspective. In G. Patrinos, W. Ansorge, & P. Danielson (Eds.), *Molecular Diagnostics* (3rd ed.). Elsevier. pp. 273-279

Abstract

Genome instability shapes the genomic landscape of cancers which takes many forms with variable clinical outcome. In cancer cytogenetics, the analysis of chromosomal copy number aberrations or the consequences thereof has routinely been assessed by FISH, PCR or IHC for candidate chromosomal locations. More recently, microarrays were introduced as a genome-wide assay. Now, developments in Next Generation Sequencing (NGS) together with informatics is causing the boundaries between cytogenetics and molecular pathology to fade, dramatically change daily cancer cytogenetic practice. The term “cytogenomics” marks this change. Information on focal chromosomal insertions and deletions, translocations, point mutations and single nucleotide polymorphisms can all be collected at once. In hereditary genome diagnostics, NGS cytogenomics has solved many rare monogenic disease and is now becoming an important tool for genetic counselling in many clinics. Even though cancer is all but a rare disease, NGS cytogenomics is less developed for cancer diagnostics, particularly in non-academic institutions. We argue that this delay can be attributed to financial, technical and analytical challenges which are unique to cancer genome analysis. First of all, the required infrastructure to perform NGS, such as sequencers and IT resources, is frequently beyond the scope and capacity of many hospitals. Additionally, in routine clinical practice, surgical resection or other invasive procedures are required to obtain tumor tissue which is routinely fixed in formalin. Finally, tumor tissue has normal cells admixed and genetic aberrations often originate from a heterogeneous tumor clone population. Each of these factors imply a more challenging NGS laboratory and data interpretation routine. NGS laboratory solutions are in active development, with a focus on sensitivity, specificity and processing time as well as a reduction in cost of cancer cytogenomic analysis. With this review we aim to provide a perspective on the current developments and the challenges awaiting us in cancer cytogenetics, based on the transition of hereditary cytogenomics in the Netherlands, and supplemented by our early NGS experiences in daily pathology practice.

Keywords: NGS, cytogenetics, cytogenomics, diagnostics

I. A Tumor Presents

When a patient presents with symptoms of cancer, there is an immediate need for a proper diagnosis, “do I indeed have a cancer?” And if it is a tumor, the next question regards consequence: “Is the tumor benign or malignant? Is it curable, or how long will I live? What are my treatment options? And how will the disease impact the quality of my life and my loved ones?” To come to a diagnosis, a step-wise process is undertaken, which can require invasive diagnostic techniques to obtain the tissue. An essential step in this process is the pathological review of the tumor material.

If the diagnosis points to a malignant tumor, the treating physician is generally joined by a multidisciplinary team, a so called ‘tumor board’, to discuss treatment options. In the Netherlands the treating physician is frequently a medical oncologist, or - depending on tumor type - a neuro-oncologist, dermatologist, pulmonologist, hematologist, pediatrician or gynecologist. At the tumor board, the case is presented and surgeons and radiologists provide information on the anatomic localization, extent of invasion and metabolic activity, while the pathologist classically provides information on the macro- and microscopic appearance of the tissue and morphology of cell and nuclear structures as well as their (immuno-) histopathological make-up. For the treatment of solid tumors, surgery still provides one of the most common and adequate remedies. Chemotherapy, including agents targeted to specific characteristics of the cancer, are also important components of the treatment palette, which is frequently supported by radiology and/or immunotherapy. The physical strength of the patient, as judged by the treating physician, is another important consideration the board will take into account when formulating treatment options, since most cancer therapies are physically taxing. The patient, advised by the treating physician, decides upon a treatment option, considering chances of successful cancer cure and the impact on the quality of life.

Until recently, apart from limited targeted mutation analysis by PCR, structural variation analysis by FISH and immunohistochemistry to assess protein overexpression, a comprehensive genomic profile of a tumor remained largely undisclosed in tumor boards. Considering that cancer is a disease of the genome [1], this molecular information is critical to understand the etiology and vulnerabilities of the cancer.

In recent years, there have been important developments in molecular techniques that bring improvements in cost effectiveness, speed, accuracy, specificity and sensitivity. The most notable development is the increase in DNA sequencing

speed, cost and throughput through the introduction of Next Generation Sequencing (NGS). These developments will soon allow routine clinical labs to sequence a genome for less than a thousand dollars [2]. Cancer genomes are sequenced with increased fidelity, resolution and detail. As a consequence, an increasing number of molecular biomarkers, which harbor clinically relevant information, are being discovered and validated. Another consequence is that targeted therapies clinically approved for one (sub-) type of cancer can be applied to other cancer types with the same targeted alteration, giving rise to the recent introduction of basket and umbrella trials [3]. In these trials, patients are grouped by molecular markers instead of disease type. Translation of these biomarkers into daily clinical practice requires adoption of NGS techniques into routine diagnostics and acquisition of requisite expertise by clinical laboratories. The ultimate goal is to provide an increasingly comprehensive personalized diagnosis for all cancer patients, by providing tumor boards with a complete clinically relevant, molecular cytogenomic picture of the tumors.

II. From microscopic examination to cytogenomics

For both cellular and molecular pathology, tumor tissue is obtained from fluids, needle biopsy or surgical resection and is transported to the pathology laboratories. The tumor is macroscopically examined dissected and routinely fixed in formalin. Formalin fixation, followed by paraffin embedding (FFPE) of tumor specimens has the advantage that it can be stored at room temperature for extended periods of time while stably preserving the (intra-) cellular morphology of the tumor. Alternative fixation procedures, also geared to maintain tissue, cellular and nuclear structures, include formic acid or flash freezing in liquid nitrogen. Sections are made from the fixed material to perform chemical or immunohistochemical staining, followed by histopathological evaluation of the tumor tissue by microscopy. Various molecular techniques have been developed that can make use of these routinely processed archival clinical specimens [4]. Storage of FFPE samples has led to the accumulation of decades' worth of patient materials. Much of the contents of these archives is accompanied by patients' history of treatment and disease course, and thereby serve as an invaluable resource for retrospective biomarker research [5], [6]. All projects investigating cancers, and particularly those with long survival, benefit from long-term clinical follow-up. For example, some patients with diffuse low-grade glioma can survive more than 25 years and access to a large series of FFPE archival material was critical to discover that distal 10q deletion is significantly less common in tumors of these long-term survivors [7]. In the Netherlands all biopsies collected in pathology

labs are recorded in a nationwide pathology registry called PALGA Pathologisch Anatomisch Landelijk Geautomatiseerd Archief [8]. PALGA enables researchers to search for samples across pathology labs throughout the country to include in their study.

Besides long-term storage, a great advantage of FFPE is that the areas of tumor tissue can be marked by a pathologist and straightforwardly scraped away from normal tissue parts prior to DNA isolation [9]. The use of FFPE tumor specimens for molecular analysis, however, also comes with drawbacks. Formalin-fixation fixes all proteins in their position, ideal for morphology, however DNA is damaged in the process. Damage to DNA extracted from FFPE specimens includes; 1) nicks and gaps, 2) modification by deamination – hydrolysis of cytosine to uracil, – 3) fragmentation, and 4) cross-linking, limiting the utility of the DNA for downstream analysis. The structure of RNA is similarly affected by the fixation process. Moreover, due to the unstable nature of mRNA its integrity can be severely compromised *ex vivo* prior to formalin fixation. Formalin fixation is furthermore detrimental for proteomics analysis. Alternatively, fresh or frozen tissue may be used for cytogenomics, but also here, precautions need to be taken to ensure preservation of nucleic acids for adequate testing.

Acquisition of the tissue material in itself also presents a number of challenges for cytogenomic diagnostics. Vital parts of the body – such as the brain – require a conservative resection strategy to obtain tissue material. Or frequently only needle biopsies are taken. These biopsy samples will have to suffice for both histopathological and molecular diagnostic procedures, resulting in variable and often very limited amounts of tissue material for DNA isolation. In addition, the tumor tissue material is frequently a heterogeneous mix of genetically distinct tumor cells with admixed normal cells, which makes the downstream NGS data analysis all the more difficult. In some cases, local sampling can cause specific tumor clones simply to be absent from the biopsy altogether. Once the tissue sample is collected, standardization and automation is not straight-forward, particularly since pathology review prior to DNA or RNA isolation is crucial [9].

Despite the discussed challenges, there are several examples of molecular assays that have successfully been implemented in clinical practice. Techniques such as the polymerase chain reaction (PCR), fluorescent *in situ* hybridization (FISH), Comparative Genome Hybridization (CGH) and array CGH assays [4], [10], have been implemented in the clinic through validation of the molecular procedures, and meticulous description of analysis workflows in standard operating procedures (SOP). NGS has also made its entrance onto the cancer diagnostic stage, particularly for the analysis of small hotspot panels consisting of 200 loci. Currently, most pathology labs in the Netherlands offer these services [11], [12].

III. The liquid biopsy promise for cancer diagnostics

In hereditary genome diagnostics blood has traditionally been the main source for DNA isolation and molecular assays. Acquisition of blood is highly standardized and uniform, which delivers a homogenous and stable source for nucleic acid extraction. The use of robotics alleviates the manual workload, while it increases the reproducibility and throughput of sample processing.

The liquid biopsy - such as blood, plasma, sputum, cervical smears or self-samples and urine - also holds great promise for cancer diagnostics. We should anticipate liquid biopsy in cancer population screening, early detection, diagnosis/prognosis, therapy selection and disease/treatment monitoring. Apart from overcoming the limitations associated with tissue biopsies, the liquid biopsy could potentially offer solutions to the described heterogeneity issue. Current assays with single molecular markers set a precedent for the success of liquid biopsies in cancer diagnostics, examples are the Human Papilloma Virus (HPV) test for cervical cancer screening on cervical smears or self-collected samples [13], the prostate specific antigen (PSA) test for prostate cancer and the fecal occult blood test (FOBT) for colorectal cancer screening.

In many academic hospitals, hereditary genome diagnostics has set the stage for liquid biopsies with the introduction of the non-invasive prenatal testing (NIPT) by NGS [14]. To perform NIPT DNA is extracted from the maternal blood which also includes a small fraction of fetal DNA. Subsequently, copy number NGS is performed to determine aneuploidy in the minute amounts of fetal DNA [15]. Diagnostic screens for cancer could very much resemble the NIPT procedure. In fact, during routine NIPT procedure, researchers at the hospital in Leuven (Belgium) incidentally observed genomic aberrations, known to occur in Hodgkin's lymphoma, and could confirm an early stage of the disease in this patient [16]. Given proper patient consent, an approach like NIPT (i.e., NGS on liquid biopsies) thus has the potential to be used for early detection or even population screening of cancer.

There is currently no general consensus as to which liquid biopsy, which method or biomolecule is most informative for cancer diagnostics. This will depend on cancer type and diagnostic goal, such that alternate liquid sources and analytical methods will be applicable. Existing tumor tissue biomarkers may be retrieved from liquid biosources such as cell free DNA (cfDNA) which is shed by the tumor in the blood circulation. Alternatively, surrogate signatures for tumor tissue biomarkers may be obtained from mRNA, i.e. extracted from blood platelets, as we have recently demonstrated at the Cancer Center Amsterdam [17]. Besides blood as a source for liquid biopsy, other non-invasive biopsy samples are also under investigation as a

source for cancer screening and monitoring; like sputum, urine, cerebrospinal fluid, cervical, anal, penile or oral swabs. The primary challenge for the clinical implementation of liquid biopsy diagnostics is of course the clinical performance, both sensitivity and specificity, of the DNA test. NGS along with novel PCR techniques - such as droplet digital PCR - currently offer the highest sensitivity to detect mutations and aberrations. DNA remnants can be very limited in the various liquid biopsies and hence a crucial step in the process is to enrich for circulating tumor DNA (ctDNA) and keep background of non-aberrant signal low. Besides these laboratory challenges, special bioinformatics procedures will have to be adapted to allow the introduction of liquid biopsy cancer cytogenomics into clinical diagnostic practice.

IV. Cytogenomic applications

Chromosomal aberrations

In cancer research we use copy number profiling to assess prognostic and predictive biomarkers in various tumor types including colon, LGG and head & neck tumors [7], [18], [19]. A routine diagnostic application in our clinic is the assessment of clonal relationship between multiple tumors from a single patient. The assay determines whether multiple tumors can be regarded as independent tumors or a single primary tumor with metastasis. The clonal relationship can be interpreted in an automated fashion by calculating a likelihood ratio to distinguish tumor pairs [20]. In our daily practice we usually complement this procedure by calculating a correlation of the aberrant segment values (figure 1), which has provided valuable insights in the clinic for effective cancer treatment [21]. Copy number aberrations are classically investigated by FISH, karyotyping of condensed chromosomal structures, CGH and later array CGH [10]. Designing probes for the FISH and CGH techniques are a time consuming and expensive undertaking and required prior knowledge on the genomic sequences of interest.

Whole genome sequencing (WGS) brings an unbiased approach as the complete DNA content of the acquired nuclei is simply being read out by NGS. An equivalent resolution to microarrays (180k features) can already be achieved with a limited amount of data where only a tenth of the (human) genome – 0,1X coverage – is actually sequenced. This low pass or shallow whole genome sequencing (sWGS) allows for cheap and efficient copy number profiling, even when using DNA isolated from FFPE tissue [6]. The general approach involves dividing the human reference genome into non-overlapping fixed sized bins. After mapping the reads to the reference genome, the raw read count is determined by the number of reads

contained in each bin. After GC content and “mappability” correction, filtering, segmentation and calling, this read count reflects an accurate copy number state of the analyzed sample.

Actionable Hot Spot Gene panel

Investigation of predictive biomarkers to assess tumor susceptibility to targeted therapies can make use of a focused approach by considering only a limited set of relevant genes to assay. For instance, melanoma patients are tested for the *BRAF* p.Val600Glu mutation. Presence of the *BRAF* p.Val600Glu mutation in melanoma indicates sensitivity to vemurafenib, an enzyme inhibitor which acts on the *BRAF/MEK/ERK* pathway to cause programmed cell death in late stage melanoma [22]. Besides *BRAF* there are currently around 50 other so called “actionable” genes. Instead of assaying all exons of these genes, a “hot-spot” approach can be implemented to target for example only around 250 of the commonly mutated loci with known clinical relevance. Currently, the turnaround time from sample to “hot-spot” result is one week in our clinical practice [11], however there is demand for even shorter turnaround times. Furthermore, we are dealing with FFPE specimens and the amount of tumor tissue is often limited, particularly for lung biopsies. Depending on the protocol used, the minimal required DNA for performing the assays ranges from 10ng to 150ng. High level amplifications, such as the

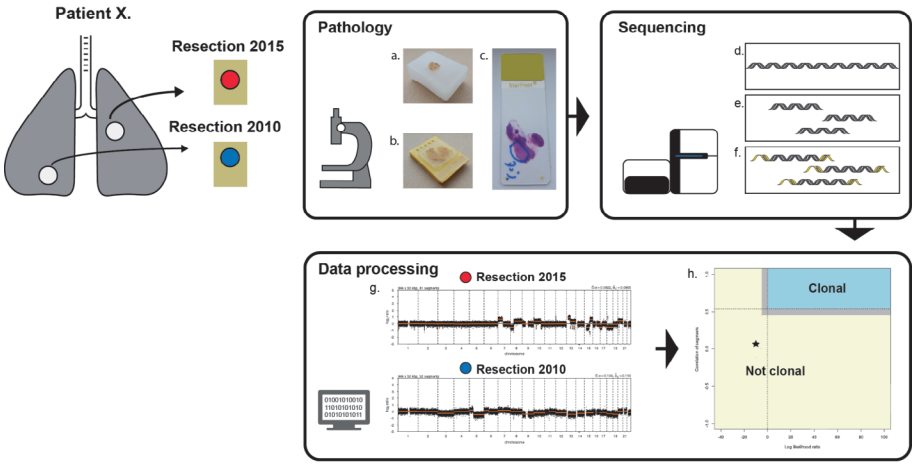

Figure 1. Workflow of clonality test. The primary resected tumor from the lung (2010) is compared to a secondary resected tumor from the lung (2015). Tissue is formalin fixed and embedded in paraffin (a, b). Sections are cut (10 micron) and HE stained followed by demarcation of tumor tissue (d). DNA is isolated from cells scraped from the marked region (d). The isolated DNA is randomly fragmented (e) and sequencing adapters are ligated for compatibility with sequencing technology (f). The resulting sequencing data are processed by QDNAseq to produce a genome wide copy number profile. Log likelihoods and correlation of segments is calculated to determine the clonality score of the two samples. The clonality score is located in the lower left quadrant which indicates the samples have no clonal relationship (h).

actionable *ERBB2 (HER2/neu)* gene can also simultaneously be detected with these panels [4], [23] by calculating deviations from the amplicon depth of coverage (DOC) over a set of control samples [24].

Exome-wide mutation analysis

In the Netherlands, only academic hospitals are allowed to perform hereditary genome diagnostics. To investigate the genetic cause of a congenital disease or cancer predisposition whole exome sequencing (WES) can be utilized. Currently, this is the most data intensive and demanding assay used in routine clinical practice at the VU University Medical Center (VUmc) in Amsterdam. If there is only a single affected patient in the family, the preferred approach is to sequence the family trio: the mother, father and the affected child. The amount of data per case can add up to 100GB and the analysis can take days to process computationally. Clinical geneticists are tasked with the annotation, review and interpretation of the reported variants, taking into account inheritance models to find the causal variant. While WES is primarily designed for qualitative analysis of point mutations, it can also be used in a quantitative way in which the copy number state of genomic regions is calculated based on the abundance of the enriched molecules of DNA. The abundance is represented by the number of reads which align to the genomic regions and needs to be corrected for a variable enrichment efficiency [25]. Not only high level amplifications, but also single copy number gains and losses may be detected, important for neuroblastomas or diffuse low grade gliomas, for example [4]. Alternatively, abundance of reads for genomic regions can also be calculated by taking into account “off-target reads” [26] which is essentially background signal and a product of the inefficiency of the enrichment protocol. The sWGS procedure described above is a more robust alternative. It requires a separate sequence run with a small aliquot of the intermediate sequence library obtained prior to the WES enrichment protocol. Since the sequence depth is limited for sWGS additional costs are marginal [6].

Translocations detection

The financial and practical feasibility of using deep WGS (> 30X coverage) in a diagnostic setting is currently being evaluated in hereditary clinical genetics laboratories. WGS will consider up to 95% of the genome - while WES considers up to 2% - and additionally reveal translocations in addition to copy number changes and mutations, and hence make the cytogenomic picture more complete. Not only for hematological malignancies, but also for an array of solid tumors, translocations have important therapeutic consequences. For example, lung

cancers that carry the EML-ALK translocation respond to the inhibitor Crizotinib [27]. Translocations or the consequences thereof are currently assayed in an elective fashion using immunohistochemistry, FISH and/or PCR techniques [4]. These techniques have serious drawbacks and suffer sensitivity, particularly since the exact translocation site may vary from tumor to tumor and genes may have different translocation partners. WGS is an alternative, but not yet within reach of the cancer cytogenomic clinic. Therefore, others and we are making efforts to develop targeted translocation NGS. Most attractive solutions include enrichment through hybrid selection panels such as WES. The BreaKmer procedure [28] makes use of WES data by focusing on the captured DNA molecules which span a translocation site. An alternative approach to translocation detection, is targeted locus amplification [29]. While also being translocation partner independent, this technique is much less demanding with regards to the exact translocation site, which may be up to 1 Mb away from the designed probes. Yet, other than BreaKmer, this approach is not yet applicable to DNA isolated from FFPE.

V. Implementation of cytogenomics in the clinic

The quality of a molecular cytogenomic diagnostic assay is crucial to determine treatment options for the patients. Predictive biomarkers are being assayed to reveal susceptibility of a tumor to a particular cytotoxin or targeted therapy. Suboptimal results can lead to selection of an inefficient therapy or can even have a detrimental effect on the wellbeing and survival of the patient. Prognostic biomarkers are being assayed to predict the (progression free) survival of the patient and form an important determinant for treatment selection. For example, an 80-year-old patient can choose to not have a tumor treated that is predicted to remain indolent for several years, whereas a 20-year-old patient would be much more likely to select (taxing) treatment when faced with a similar prognosis.

When it comes to the introduction of cytogenomic assays into the clinic, the hereditary genome diagnostic laboratories in the Netherlands have experienced a number of development cycles of novel molecular techniques that are adopted and adapted from research to a clinical application. Diagnostic laboratories adhere to a specific set of requirements to warrant quality, validity, reliability and reproducibility of a molecular assays. The molecular technique needs refinement and standardization to comply to diagnostic standards before it can be used in routine diagnostics [30]. Clinical molecular specialists either in hereditary genome diagnostics or molecular pathology are tasked to interpret the molecular data and report the clinically relevant information back to the pathologist, treating physician or tumor board. At the same time, they need to be aware of the latest

developments in molecular biology techniques and have an overview of the existing needs in the clinic which is essential to continue innovation. The clinical molecular specialist will also need a level of understanding of bioinformatics and IT infrastructure issues, as the data analysis and interpretation become increasingly complicated.

VI. Bioinformatics & Data analysis

The growing amount of data brought by NGS will make research and clinical applications more and more reliant on a robust IT infrastructure and organization. Current academic hospital IT organizations are mainly geared towards supporting office applications and delivering robust information systems to securely manage the flow of patient information. However, often these organizations lack the ability to quickly respond to developments in the research setting and they often lack generic high performance computing systems for researchers to process their data. Additionally, the large differences in business culture between hospital and research are often a source of frustration and miss communication. It is very difficult for a single person to oversee all these disciplines, therefore a specialized team is required to bridge the gap in the genomics era between an ever changing research environment, fast paced genome diagnostics, and a generic hospital IT setting.

Most peripheral Pathology laboratories in The Netherlands currently do not have sufficient IT resources to handle the large data flow generated by NGS. Such labs are dependent on turnkey solutions provided by commercial companies such as Sofia Genetics, CLC Bio, DNA star, JSI- medical systems or DNAnexus. These companies provide software to analyze small-scale sequencing applications which focus on a limited set of loci such as hot-spot panels or gene panels. A graphical user interface is provided to allow laboratory technicians and specialists to analyze and interpret the data. The investment required to acquire a commercial product is relatively small and alleviates the need for compute and storage infrastructure. Commercial companies provide an attractive and efficient means of adopting well established analysis workflows in the clinic but are limited in their ability to co-develop new analysis workflows coming from research labs. Most of these software packages are locally installed tools and lack the capability to scale up when additional resources are needed. To address this problem, Cartagenia, DNAnexus and Sofia Genetics provide their software as a service (SaaS) and use the scaling capabilities of public cloud compute providers to provide the required resources. This solution, however, requires strict compliance with security and privacy regulations, as patient data are housed outside of hospital networks.

VII. Concluding remarks

Even though clinical genetics and molecular pathology are distinctly different disciplines, their common interest in cytogenomics calls for extensive collaborations. The laboratory of tumor genetics (LTG) at the Radboud University Medical Center Nijmegen in The Netherlands has pioneered such collaboration. They focus both on germline mutations to determine predisposition to cancer as well as on acquired somatic mutations to determine prognostic and/or predictive outcome to treatment regimens. Their success has inspired us as well as many other academic hospitals in The Netherlands to follow in their footsteps. Sharing allows efficient exploitation of lab equipment, computing and storage infrastructure by both diagnostic and research applications for multiple departments. Moreover, colocation of NGS and bioinformatics on the one hand and research and diagnostic expertise on the other strengthens collaboration and shortens the said development cycle.

Confronting the new demands of cytogenomic diagnostics is the current economic climate. In the Netherlands health care providers are encouraged to adopt a commercial and competitive model which makes them independent of government financial support. In many cases this leads to a lack of high risk investment such as acquisition of cutting edge technology and IT infrastructure. This situation has a negative effect on the development cycle from research to clinical application. Nevertheless, cancer cytogenomics is a moving train on a track paved by hereditary cytogenomics. With the 1000-dollar genome within reach, the penetrance of cancer NGS in daily clinical routine will only increase.

References

[1] D. Hanahan and R. A. Weinberg, “Hallmarks of cancer: The next generation,” *Cell*, vol. 144, no. 5. pp. 646–674, 2011.

[2] P. Muir, S. Li, S. Lou, D. Wang, D. J. Spakowicz, L. Salichos, J. Zhang, G. M. Weinstock, F. Isaacs, J. Rozowsky, and M. Gerstein, “The real cost of sequencing: scaling computation to keep pace with data generation.,” *Genome Biol.*, vol. 17, no. 1, p. 53, 2016.

[3] D. M. Hyman, I. Puzanov, V. Subbiah, J. E. Faris, I. Chau, J.-Y. Blay, J. Wolf, N. S. Raje, E. L. Diamond, A. Hollebecque, R. Gervais, M. E. Elez-Fernandez, A. Italiano, R.-D. Hofheinz, M. Hidalgo, E. Chan, M. Schuler, S. F. Lasserre, M. Makrutzki, F. Sirzen, M. L. Veronese, J. Tabernero, and J. Baselga, “Vemurafenib in Multiple Nonmelanoma Cancers with BRAF V600 Mutations.,” *N. Engl. J. Med.*, vol. 373, no. 8, pp. 726–36, 2015.

[4] R. J. Hastings, N. Bown, M. G. Tibiletti, M. Debiec-Rychter, R. Vanni, B. Espinet, N. van Roy, P. Roberts, E. van den Berg-de-Ruiter, A. Bernheim, J. Schoumans, S. Chatters, Z. Zemanova, M. Stevens-Kroef, A. Simons, S. Heim, M. Salido, B. Ylstra, D. R. Betts, Tumour Best Practice meeting, and Eurogentest, “Guidelines for cytogenetic investigations in tumours.,” *Eur. J. Hum. Genet.*, vol. 24, no. 1, pp. 6–13, Jan. 2016.

[5] N. Blow, “Tissue preparation: Tissue issues.,” *Nature*, vol. 448, no. 7156, pp. 959–963, 2007.

[6] I. Scheinin, D. Sie, H. Bengtsson, M. A. van de Wiel, a. B. Olshen, H. F. van Thuijl, H. F. van Essen, P. P. Eijk, F. Rustenburg, G. A. Meijer, J. C. Reijneveld, P. Wesseling, D. Pinkel, D. G. Albertson, and B. Ylstra, “DNA copy number analysis of fresh and formalin-fixed specimens by shallow whole-genome sequencing with identification and exclusion of problematic regions in the genome assembly,” *Genome Res.*, vol. 24, no. 12, pp. 1–11, Dec. 2014.

[7] H. F. van Thuijl, I. Scheinin, D. Sie, A. Alentorn, H. F. van Essen, M. Cordes, R. Fleischeuer, A. M. Gijtenbeek, G. Beute, W. A. van den Brink, G. A. Meijer, M. Havenith, A. Idbaih, K. Hoang-Xuan, K. Mokhtari, R. G. Verhaak, P. van der Valk, M. A. van de Wiel, J. J. Heimans, E. Aronica, J. C. Reijneveld, P. Wesseling, and B. Ylstra, “Spatial and temporal evolution of distal 10q deletion, a prognostically unfavorable event in diffuse low-grade gliomas.,” *Genome Biol.*, vol. 15, no. 9, p. 471, Sep. 2014.

[8] M. Casparie, A. T. M. G. Tiebosch, G. Burger, H. Blauwgeers, A van de Pol, J. H. J. M. van Krieken, and G. A. Meijer, “Pathology databanking and biobanking in The Netherlands, a central role for PALGA, the nationwide histopathology and cytopathology data network and archive.,” *Cell. Oncol.*, vol. 29, no. 1, pp. 19–24, 2007.

[9] H. F. Van Essen and B. Ylstra, “High-resolution copy number profiling by array CGH using DNA isolated from formalin-fixed, paraffin-embedded tissues,” *Methods Mol. Biol.*, vol. 838, pp. 329–341, 2012.

[10] B. Ylstra, P. van den IJssel, B. Carvalho, R. H. Brakenhoff, and G. A. Meijer, “BAC to the future! Or oligonucleotides: A perspective for micro array comparative genomic hybridization (array CGH),” *Nucleic Acids Research*, vol. 34, no. 2. pp. 445–450, 2006.

[11] D. Sie, P. J. F. Snijders, G. A. Meijer, M. W. Doeleman, M. I. H. Van Moorsel, H. F. Van Essen, P. P. Eijk, K. Grünberg, N. C. T. Van Grieken, E. Thunnissen, H. M. Verheul, E. F. Smit, B. Ylstra, and D. A. M. Heideman, “Performance of amplicon-based next generation DNA sequencing for diagnostic gene mutation profiling in oncopathology,” *Cell. Oncol.*, vol. 37, no. 5, pp. 353–361, Oct. 2014.

[12] A. Boleij, B. B. J. Tops, P. D. M. Rombout, E. M. Dequeker, M. J. L. Ligtenberg, J. H. van Krieken, and Dutch RAS EQA Initiative, “RAS testing in metastatic colorectal cancer: excellent reproducibility amongst 17 Dutch pathology centers.,” *Oncotarget*, vol. 6, no. 17, pp. 15681–9, Jun. 2015.

[13] M. Schiffman, N. Wentzensen, S. Wacholder, W. Kinney, J. C. Gage, and P. E. Castle, “Human papillomavirus testing in the prevention of cervical cancer.,” *J. Natl. Cancer Inst.*, vol. 103, no. 5, pp. 368–83, Mar. 2011.

[14] S. Tamminga, M. van Maarle, L. Henneman, C. B. M. Oudejans, M. C. Cornel, and E. A. Sistermans, “Maternal Plasma DNA and RNA Sequencing for Prenatal Testing,” in *Advances in clinical chemistry*, vol. 74, 2016, pp. 63–102.

[15] R. Straver, E. A. Sistermans, H. Holstege, A. Visser, C. B. M. Oudejans, and M. J. T. Reinders, “WISECONDOR: detection of fetal aberrations from shallow sequencing maternal plasma based on a within-sample comparison scheme.,” *Nucleic Acids Res.*, vol. 42, no. 5, p. e31, 2014.

[16] F. Amant, M. Verheecke, I. Wlodarska, L. Dehaspe, P. Brady, N. Brison, K. Van Den Bogaert, D. Dierickx, V. Vandecaveye, T. Tousseyn, P. Moerman, A. Vanderstichele, I. Vergote, P. Neven, P. Berteloot, K. Putseys, L. Danneels, P. Vandenberghe, E. Legius, and J. R. Vermeesch, “Presymptomatic Identification of Cancers in Pregnant Women During Noninvasive Prenatal Testing.,” *JAMA Oncol.*, vol. 1, no. 6, pp. 814–9, 2015.

[17] M. G. Best, N. Sol, I. Kooi, J. Tannous, B. A. Westerman, F. Rustenburg, P. Schellen, H. Verschueren, E. Post, J. Koster, B. Ylstra, N. Ameziane, J. Dorsman, E. F. Smit, H. M. Verheul, D. P. Noske, J. C. Reijneveld, R. J. A. Nilsson, B. A. Tannous, P. Wesseling, and T. Wurdinger, “RNA-Seq of Tumor-Educated Platelets Enables Blood-Based Pan-Cancer, Multiclass, and Molecular Pathway Cancer Diagnostics,” *Cancer Cell*, vol. 28, no. 5, pp. 666–676, 2015.

[18] J. C. Haan, M. Labots, C. Rausch, M. Koopman, J. Tol, L. J. M. Mekenkamp, M. A. van de Wiel, D. Israeli, H. F. van Essen, N. C. T. van Grieken, Q. J. M. Voorham, L. J. W. Bosch, X. Qu, O. Kabbarah, H. M. W. Verheul, I. D. Nagtegaal, C. J. a. Punt, B. Ylstra, and G. a. Meijer, “Genomic landscape of metastatic colorectal cancer,” *Nat. Commun.*, vol. 5, p. 5457, 2014.

[19] S. J. Smeets, R. H. Brakenhoff, B. Ylstra, W. N. Van Wieringen, M. A. Van De Wiel, C. R. Leemans, and B. J. M. Braakhuis, “Genetic classification of oral and oropharyngeal carcinomas identifies subgroups with a different prognosis,” *Cell. Oncol.*, vol. 31, no. 4, pp. 291–300, 2009.

[20] I. Ostrovnaya, A. B. Olshen, V. E. Seshan, I. Orlow, D. G. Albertson, and C. B. Begg, “A metastasis or a second independent cancer? Evaluating the clonal origin of tumors using array copy number data.,” *Stat. Med.*, vol. 29, no. 15, pp. 1608–21, Jul. 2010.

[21] J. L. Kuiper, M. I. Ronden, A. Becker, D. A. M. Heideman, P. van Hengel, B. Ylstra, E. Thunnissen, and E. F. Smit, “Transformation to a squamous cell carcinoma phenotype of an *EGFR* -mutated NSCLC patient after treatment with an EGFR-tyrosine kinase inhibitor,” *J. Clin. Pathol.*, vol. 68, no. 4, pp. 320–321, Apr. 2015.

[22] G. Bollag, J. Tsai, J. Zhang, C. Zhang, P. Ibrahim, K. Nolop, and P. Hirth, “Vemurafenib: the first drug approved for BRAF-mutant cancer.,” *Nat. Rev. Drug Discov.*, vol. 11, no. 11, pp. 873–86, Nov. 2012.

[23] A. V Uzilov, W. Ding, M. Y. Fink, Y. Antipin, A. S. Brohl, C. Davis, C. Y. Lau, C. Pandya, H. Shah, Y. Kasai, J. Powell, M. Micchelli, R. Castellanos, Z. Zhang, M. Linderman, Y. Kinoshita, M. Zweig, K. Raustad, K. Cheung, D. Castillo, M. Wooten, I. Bourzgui, L. C. Newman, G. Deikus, B. Mathew, J. Zhu, B. S. Glicksberg, A. S. Moe, J. Liao, L. Edelmann, J. T. Dudley, R. G. Maki, A. Kasarskis, R. F. Holcombe, M. Mahajan, K. Hao, B. Reva, J. Longtine, D. Starcevic, R. Sebra, M. J. Donovan, S. Li, E. E. Schadt, and R. Chen, “Development and clinical application of an integrative genomic approach to personalized cancer therapy.,” *Genome Med.*, vol. 8, no. 1, p. 62, 2016.

[24] M. Hoogstraat, J. W. J. Hinrichs, N. J. M. Besselink, J. H. Radersma-van Loon, C. M. A. de Voijs, T. Peeters, I. J. Nijman, R. A. de Weger, E. E. Voest, S. M. Willems, E. Cuppen, and M. J. Koudijs, “Simultaneous detection of clinically relevant mutations and amplifications for routine cancer pathology.,” *J. Mol. Diagn.*, vol. 17, no. 1, pp. 10–8, Jan. 2015.

[25] A. Magi, L. Tattini, I. Cifola, R. D’Aurizio, M. Benelli, E. Mangano, C. Battaglia, E. Bonora, A. Kurg, M. Seri, P. Magini, B. Giusti, G. Romeo, T. Pippucci, G. De Bellis, R. Abbate, and G. F. Gensini, “EXCAVATOR: detecting copy number variants from whole-exome sequencing data.,” *Genome Biol.*, vol. 14, no. 10, p. R120, 2013.

[26] T. Kuilman, A. Velds, K. Kemper, M. Ranzani, L. Bombardelli, M. Hoogstraat, E. Nevedomskaya, G. Xu, J. de Ruiter, M. P. Lolkema, B. Ylstra, J. Jonkers, S. Rottenberg, L. F. Wessels, D. J. Adams, D. S. Peeper, and O. Krijgsman, "CopywriteR: DNA copy number detection from off-target sequence data," *Genome Biol.*, vol. 16, no. 1, p. 49, Feb. 2015.

[27] E. L. Kwak, Y.-J. Bang, D. R. Camidge, A. T. Shaw, B. Solomon, R. G. Maki, S.-H. I. Ou, B. J. Dezube, P. A. Jänne, D. B. Costa, M. Varella-Garcia, W.-H. Kim, T. J. Lynch, P. Fidias, H. Stubbs, J. A. Engelman, L. V Sequist, W. Tan, L. Gandhi, M. Mino-Kenudson, G. C. Wei, S. M. Shreeve, M. J. Ratain, J. Settleman, J. G. Christensen, D. A. Haber, K. Wilner, R. Salgia, G. I. Shapiro, J. W. Clark, and A. J. Iafrate, "Anaplastic lymphoma kinase inhibition in non-small-cell lung cancer.," *N. Engl. J. Med.*, vol. 363, no. 18, pp. 1693–703, 2010.

[28] R. P. Abo, M. Ducar, E. P. Garcia, A. R. Thorner, V. Rojas-Rudilla, L. Lin, L. M. Sholl, W. C. Hahn, M. Meyerson, N. I. Lindeman, P. Van Hummelen, and L. E. MacConaill, "BreaKmer: detection of structural variation in targeted massively parallel sequencing data using kmers.," *Nucleic Acids Res.*, vol. 43, no. 3, p. e19, 2015.

[29] P. J. P. de Vree, E. de Wit, M. Yilmaz, M. van de Heijning, P. Klous, M. J. A. M. Verstegen, Y. Wan, H. Teunissen, P. H. L. Krijger, G. Geeven, P. P. Eijk, D. Sie, B. Ylstra, L. O. M. Hulsman, M. F. van Dooren, L. J. C. M. van Zutven, A. van den Ouweland, S. Verbeek, K. W. van Dijk, M. Cornelissen, A. T. Das, B. Berkhout, B. Sikkema-Raddatz, E. van den Berg, P. van der Vlies, D. Weening, J. T. den Dunnen, M. Matusiak, M. Lamkanfi, M. J. L. Ligtenberg, P. ter Brugge, J. Jonkers, J. A. Foekens, J. W. Martens, R. van der Luijt, H. K. P. van Amstel, M. van Min, E. Splinter, and W. de Laat, "Targeted sequencing by proximity ligation for comprehensive variant detection and local haplotyping.," *Nat. Biotechnol.*, vol. 32, no. 10, pp. 1019–25, Oct. 2014.

[30] M. M. Weiss, B. Van der Zwaag, J. D. H. Jongbloed, M. j. Vogel, H. T. Brüggenwirth, R. H. Lekanne Deprez, O. Mook, C. A. l. Ruivenkamp, M. A. van Slegtenhorst, A. Van den Wijngaard, Q. Waisfisz, M. R. Nelen, and N. van der Stoep, "Best Practice Guidelines for the Use of Next Generation Sequencing (NGS) Applications in Genome Diagnostics: A National Collaborative Study of Dutch Genome Diagnostic Laboratories," *Hum. Mutat.*, vol. 34, no. 10, p. n/a–n/a, 2013.